

# Social Decisions and Fairness Change When People's Interests Are Represented by Autonomous Agents

JAAMAS Track

Celso M. de Melo

US Army Research Laboratory, Playa Vista, CA 90094, United States  
celso.m.demelo.civ@mail.mil

Stacy Marsella

College of Computer and Information Science, Northeastern University, Boston, MA 02115, USA  
stacymarsella@gmail.com

Jonathan Gratch

USC Institute for Creative Technologies, Playa Vista, CA 90094-2536, United States  
gratch@usc.ict.edu

## ABSTRACT

Recent times have seen an emergence of a new breed of intelligent machines that act autonomously on our behalf, such as autonomous vehicles, drones, personal assistants, etc. These machines introduce a new interaction paradigm where people instruct, or program, these agents to act on their behalf with others. Here we show that this act of programming changes the way people think about the situation, often leading them to adopt a broader perspective and act more fairly. We present four studies where participants made fairer decisions in ultimatum and negotiation tasks when engaging through an agent representative, when compared to direct interaction with others. These findings emphasize the importance of understanding the cognitive factors underlying people's decision making when designing autonomous machines, if we wish to promote a fairer society.

## KEYWORDS

Agent Representatives, Decision making, Fairness

### ACM Reference format:

C. de Melo, S. Marsella, and J. Gratch. 2018. Social Decisions and Fairness Change When People's Interests Are Represented by Autonomous Agents. In *Proc. of the 17th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2018)*, M. Dastani, G. Sukthankar, E. André, S. Koenig (eds.), July 10-15, 2018, Stockholm, Sweden.

## 1 INTRODUCTION

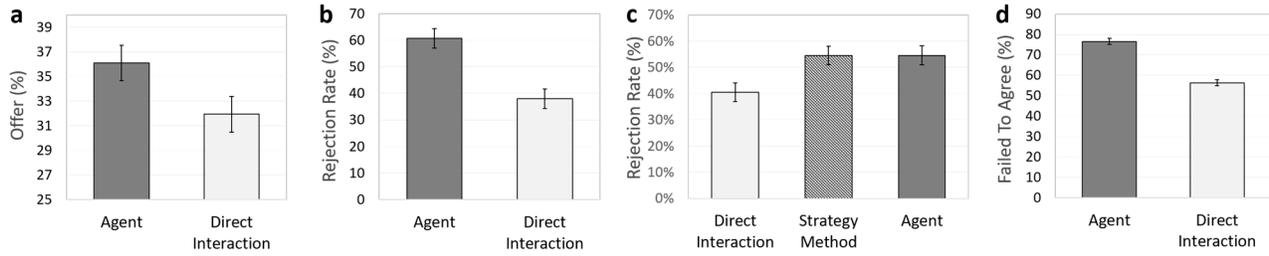
There has been increasing interest in intelligent machines – such as autonomous vehicles, drones, personal assistants, etc. – that can make autonomous decisions on people's behalf. In this paper, we ask a simple question: Do people's decisions change when they “program” an agent to act on their behalf? It is important to clarify that “programming” here is meant in the more general sense of providing high-level instructions, rather than the technical aspects of implementing those instructions. Other researchers had

noted that programming could lead people to make different decisions than they would if they were deciding “in the moment” [1]-[4]. We motivate two competing hypotheses. On the one hand, increasing the psychological distance to others – by manipulating social, temporal, and physical distance – can lead people to treat distant others less fairly [5]-[9]. Since acting with others through an agent increases the distance to others, one competing hypothesis is that: People will show reduced fairness when tasking agents to act on their behalf than when interacting directly with others (Hypothesis 1). On the other hand, another line of research suggests that making a decision ahead of time – as happens when programming – reinforces social norms, such as fairness [10]-[12]. Whereas real-time interactions require people to respond to a specific and immediate situation, programming requires the programmer to deliberate on all possible situations that might arise and to devise rules that consistently hold across all of these eventualities. This encourages the decision maker to adopt a broader perspective, consider the counterpart's position, and act more fairly. The second competing hypothesis is thus: People will show increased fairness when tasking agents to act on their behalf than when interacting directly with others (Hypothesis 2).

## 2 EXPERIMENTS

To test these hypotheses, we ran four experiments where participants engaged in decision making tasks, either through an agent representative or directly with others. In Experiment 1, participants engaged in the ultimatum and impunity games. In the ultimatum game [13], there are two players: a proposer and a responder. The proposer is given an initial endowment of money and has to decide how much to offer to the responder. Then, the responder has to make a decision: if the offer is accepted, both players get the proposed allocation; if the offer is rejected, however, no one gets anything. The standard rational prediction is that the proposer should offer the minimum non-zero amount, as the responder will always prefer something to nothing. In practice, people usually offer 40 to 50 percent of the initial endowment and low offers (about 20 percent) are usually rejected [14]. This behavior is usually explained by a concern with fairness and a fear of being rejected [15]. The impunity game is similar, but only allows a symbolic rejection – i.e., if the responder rejects

Proc. of the 17th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2018), M. Dastani, G. Sukthankar, E. André, S. Koenig (eds.), July 10-15, 2018, Stockholm, Sweden. © 2018 International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org). All rights reserved.



**Figure 1: Experimental results:** a) Participants offered more when interacting through an agent; b) Participants were less likely to accept unfair offers when interacting through an agent; c) Acting through an agent is similar to deciding ahead of time (strategy method); d) Participants were less likely to reach an agreement with unfair negotiators.

the offer, the proposer still gets his or her share of the offer. The impunity game is, therefore, a version of the ultimatum game where responders are given less power over the outcome. Experimental results with this game show that proposers tend to offer less than in the ultimatum game, though still above the rational prediction of zero [16].

In Experiment 1, participants engaged in these games as proposers and, in support of Hypothesis 2, the results showed that they made fairer offers when acting through an agent than when interacting directly with their counterparts (Figure 1-a). In Experiment 2, participants engaged once again in these games, but this time in the role of responders. The proposers, which were scripted, always made unfair offers. Reinforcing Hypothesis 2, the results showed that participants were more likely to reject unfair offers when interacting through agent representatives than when interacting directly (Figure 1-c).

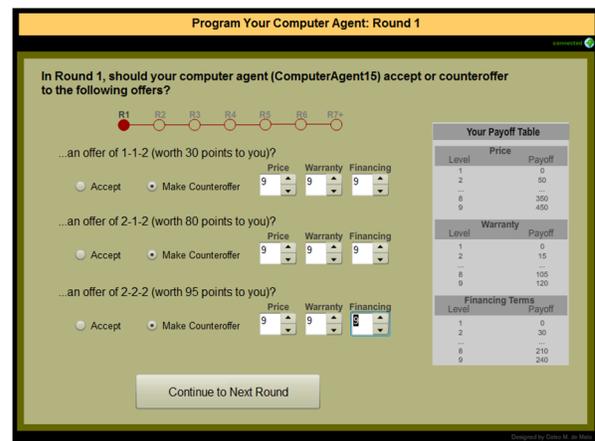
In a third experiment, we introduced a new condition where participants were asked to think about all the possible outcomes of the game ahead of time and self-report their decision before interacting with their counterparts. Notice that, in this condition, no agents are involved. In experimental economics, this procedure is usually referred to as the strategy method [11], [12]. The results showed that participants were less likely to accept unfair offers both in the strategy method and when interacting through agents, than when interacting directly (Figure 1-c). This suggests that the agent effect is driven by the fact that the decision is made ahead of time, rather than on a moment-by-moment basis as in direct interaction.

Finally, in Experiment 4 participants engaged in a typical multi-round multi-issue bargaining task with a counterpart that consistently made unfair offers. In the agent representative condition, participants had to program their agents to respond – either accept or make a counter-offer – to several possible (fair and unfair) offers (Figure 2). In support of Hypothesis 2, the results revealed that participants were less likely to reach an agreement when interacting through an agent than when interacting directly (Figure 1-d).

### 3 GENERAL DISCUSSION

At a time when autonomous agents are becoming more pervasive in society, it is important we understand if our social decisions are

changing. Here we show consistent evidence, across four experiments, that programming an agent to act on people’s behalf can change the way people think about the situation and promote the adoption of a broader perspective that leads to fairer behavior, when compared to direct interaction. This effect occurred in simple tasks (ultimatum game), tasks where the decision maker had little power (impunity game), and in more complex tasks (multi-issue bargaining). Our results also suggest that this effect is not necessarily specific to autonomous machines as participants behaved more fairly when reporting their decisions ahead of time (strategy method), even when no agents were involved. Nevertheless, autonomous agents introduce a unique opportunity to promote this reflection and deliberation in users and, consequently, encourage a fairer society.



**Figure 2: Participants programmed their agents to negotiate on their behalf by providing example decisions to different sets of offers.**

### ACKNOWLEDGMENTS

This work is supported by the National Science Foundation, under grant BCS-1419621, and the Air Force Office of Scientific Research, under grant FA9550-14-1-0364. The content does not necessarily reflect the position or the policy of any Government, and no official endorsement should be inferred.

## REFERENCES

- [1] Lin, R., Kraus, S., Oshrat, Y., & Gal, Y. (2010). Facilitating the evaluation of automated negotiators using peer designed agents. In *Proceedings of the 24th AAAI Conference on Artificial Intelligence (AAAI'10)*.
- [2] Chalamish, M., Sarne, D., & Lin, R. (2013). Enhancing parking simulations using peer-designed agents. *IEEE Transactions on Intelligent Transportation Systems*, 14(1), 492-498.
- [3] Grosz, B., Kraus, S., & Talman, S. (2004). The influence of social dependencies on decision-making: initial investigations with a new game. In *Proceedings of the International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS'04)*.
- [4] Elmalech, A., & Sarne, D. (2013). Evaluating the applicability of peer-designed agents for mechanism evaluation. *Web Intelligence and Agent Systems: An International Journal*, 12(2), 171-191.
- [5] Hoffman, E., McCabe, K., & Smith, V. (1996). Social distance and other-regarding behavior in dictator games. *American Economic Review*, 86, 653-660.
- [6] Goeree, J., McConnell, M., Mitchell, T., Tromp, T., & Yariv, L. (2010). The 1/d law of giving. *American Economic Journal: Microeconomics*, 2, 183-203.
- [7] Yu, R., Hu, P., & Zhang, P. (2015). Social distance and anonymity modulate fairness consideration: An ERP study. *Science Reports*, 5, 1-12.
- [8] Pronin, E., Olivola, C., & Kennedy, K. (2008). Doing unto future selves as you would do unto others: Psychological distance and decision making. *Personality and Social Psychology Bulletin*, 34, 224-236.
- [9] Nowak, M., & May, R. (1992). Evolutionary games and spatial chaos. *Nature*, 359, 826-829.
- [10] Rauhut, H., & Winter, F. (2010). A sociological perspective on measuring social norms by means of strategy method experiments. *Social Science Research*, 39, 1181-1194.
- [11] Güth, W., & Tietz, R. (1990). Ultimatum bargaining behavior: A survey and comparison of experimental results. *Journal of Economic Psychology*, 11, 417-449.
- [12] Oosterbeek, H., Sloof, R., & Van de Kuilen, G. (2004). Cultural differences in ultimatum game experiments: Evidence from a meta-analysis. *Experimental Economics*, 7, 171-188.
- [13] Güth, W., Schmittberger, R., & Schwarze, B. (1982). An experimental analysis of ultimatum bargaining. *Journal of Economic Behavior and Organization*, 3, 367-388.
- [14] Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., et al. (2001). In search of homo economicus: behavioral experiments in 15 small-scale societies. *American Economic Review*, 91, 73-78.
- [15] Camerer, C., & Thaler, R. (1995). Ultimatums, dictators, and manners. *Journal of Economic Perspectives*, 9, 209-219.
- [16] Yamagishi, T., Horita, H., Shinada, M., Tanida, S., & Cook, K. (2009). The private rejection of unfair offers and emotional commitment. *Proceedings of the National Academy of Science USA*, 106, 11520-11523.